Analysis of left-censored longitudinal data with application to viral load in HIV infection

HÉLÈNE JACQMIN-GADDA*, RODOLPHE THIÉBAUT
Institut National de la Santé et de la Recherche Médicale U330, 146 rue Léo Saignat,
33076 Bordeaux cedex, France
helene.jacqmin-gadda@bordeaux.inserm.fr

GENEVIÈVE CHÊNE
Département d’Informatique Médicale, Université Victor Ségalen Bordeaux II, 146 rue Léo Saignat,
33076 Bordeaux cedex, France

DANIEL COMMENGES
Institut National de la Santé et de la Recherche Médicale U330, 146 rue Léo Saignat,
33076 Bordeaux cedex, France

SUMMARY

The classical model for the analysis of progression of markers in HIV-infected patients is the mixed effects linear model. However, longitudinal studies of viral load are complicated by left censoring of the measures due to a lower quantification limit. We propose a full likelihood approach to estimate parameters from the linear mixed effects model for left-censored Gaussian data. For each subject, the contribution to the likelihood is the product of the density for the vector of the completely observed outcome and of the conditional distribution function of the vector of the censored outcome, given the observed outcomes. Values of the distribution function were computed by numerical integration. The maximization is performed by a combination of the Simplex algorithm and the Marquardt algorithm. Subject-specific deviations and random effects are estimated by modified empirical Bayes replacing censored measures by their conditional expectations given the data. A simulation study showed that the proposed estimators are less biased than those obtained by imputing the quantification limit to censored data. Moreover, for models with complex covariance structures, they are less biased than Monte Carlo expectation maximization (MCEM) estimators developed by Hughes (1999). The method was then applied to the data of the ALBI-ANRS 070 clinical trial for which HIV-1 RNA levels were measured with an ultrasensitive assay (quantification limit 50 copies/ml). Using the proposed method, estimates obtained with data artificially censored at 500 copies/ml were close to those obtained with the real data set.

Keywords: Left-censoring; Linear mixed effects model; Repeated measures.

1. INTRODUCTION

HIV-1 RNA level or viral load is a widespread marker of the evolution of HIV-1-infected patients (Albert et al., 1998). It has been recognized as the best prognostic marker with CD4 + cell counts

*To whom correspondence should be addressed

© Oxford University Press (2000)
(Mellors et al., 1996) and the reduction in HIV-1 RNA level is frequently used as the primary endpoint in clinical trials of anti-retroviral therapy. However, the major problem with the HIV-1 RNA level is due to measurement limitations. Indeed, all available assays today have a lower limit of quantification, generally between 500 and 20 copies/ml, so the measures are left-censored. Moreover, with highly active anti-retroviral treatments, the proportion of subjects with an RNA level below the limit has increased.

In clinical trials, two methods are commonly used to deal with this problem. The outcome may be defined as the percentage of patients having an HIV-1 RNA level below the quantification limit in each group at the end of the study. Providing that there is no loss of follow-up, this analysis is unbiased but leads to a loss of information. More importantly, the threshold depends on the assay used and may not be clinically relevant. The other strategy is to define the endpoint as the mean difference of HIV-1 RNA level between baseline and the end of the study. In this case, the quantification limit is imputed to all censored values leading to a conservative test. The use of survival analysis methods has been proposed to compare the reduction of HIV-1 RNA level between groups by taking into account the censoring (Marschner et al., 1999). In fact, if the baseline RNA level is completely observed and the RNA level at the end of the trial is left-censored, the reduction is right-censored, as is classically described for survival time. Thus, if the assumption of non-informative censoring is valid, the cumulative frequency distribution and the median reduction may be computed using the Kaplan–Meier approach, and the log-rank test may be used to compare treatment groups. When censoring is informative, or for adjustment on covariates, a parametric alternative is to use a censored linear regression analysis. Such methods lead to an unbiased test for the reduction of RNA level but do not describe the evolution of RNA level with time.

The classical model for the analysis of longitudinal Gaussian data is the linear mixed effects model (Laird and Ware, 1982), which is commonly used to study the progression of HIV infection markers (Boscardin et al., 1998). When the marker is the viral load, the quantification limit (Keet et al., 1997) or half of this limit (O’Brien et al., 1998) may be imputed for censored measures. However, the bias for such a strategy may be large. Paxton et al. (1997) used a two-stage imputation approach. They fitted a mixed-effects model using half of the limit; then, using the results of this analysis, they imputed new values and refitted the model. They found little difference in the estimates of fixed effects between the two samples but estimates of variance components were changed. None of these methods takes the imputation process into account in order to compute the variance of the parameter estimates.

More recently Hughes (1999) proposed an MCEM algorithm to obtain maximum likelihood estimates of parameters of the mixed effects linear model when data are left- or right-censored. As the expectation step of the EM algorithm is intractable, the author used a Gibbs sampler.

The aim of the present paper is also to present a method to obtain maximum likelihood estimates of parameters of the mixed effects linear model when data are left-censored. However, we propose a direct maximization of the likelihood without using the EM algorithm or a Monte Carlo method. This approach allows us to estimate models with a more complex covariance structure. Moreover, we have developed formulae to estimate subject-specific deviations and individual predictions of the outcome, and a method to check assumptions of the model. This article is completed by: a simulation study to compare our method with the MCEM algorithm and with naive methods; and a validation study using a real data set. The linear mixed effects model is described in Section 2. We develop the likelihood when data are left-censored in Section 3. Computational details for the maximum likelihood estimation are given in Section 4. The formulae to obtain subject-specific deviations and individual estimations for the outcome are derived in Section 5. Sections 6 and 7 present the simulation study and the application.
2. The Model

We used a linear model for longitudinal Gaussian data with a general covariance structure including a random component, a Gaussian stochastic process and an observational error (Laird and Ware, 1982; Diggle, 1988; Jones and Boati-Boateng, 1991). This covariance structure is more general than that used by Hughes (1999), which did not include a Gaussian stochastic process. Let $Y_i$ denote the $n_i$-vector of outcomes for subject $i$, $i = 1, \ldots, N$, measured at times $t_i = [t_{i1}, \ldots, t_{im_i}]^T$. The number and the times of observations may be variable for each subject. The model is defined by:

$$Y_i = X_i \beta + Z_i \gamma_i + w_i(t_i) + \epsilon_i$$  \hspace{1cm} (1)

where $\beta$ is a $p$-vector of fixed effects and $X_i$ is the corresponding $n_i \times p$ design matrix; $\gamma_i$ is a $q$-vector of random effects and $Z_i$ is an $n_i \times q$ sub-matrix of $X_i$. The vectors $\gamma_i$ are assumed to be normally distributed with mean 0 and covariance matrix $B$. Typically, the matrix $X_i$ includes a vector of one (1), the vector $t_i$ and possibly powers of $t_i$, and fixed or time-dependent covariates. For a model with random intercept and slope, $Z_i$ is the matrix $[1 \ t_i]$.

The $\{w_i(t): t \in \mathbb{R}\}$ are independent Gaussian processes. So $w_i(t_i)$ is multivariate Gaussian with mean zero and covariance matrix $\Sigma_i$. Some authors propose using a continuous-time autoregressive process (Diggle, 1988; Jones and Boati-Boateng, 1991), while others assume $w_i(t_i)$ to be a Brownian motion (Taylor et al., 1994; Boscardin et al., 1998). In these cases, the elements of the matrix $\Sigma_i$ are respectively $\sigma_w^2 \exp(-\alpha |t_i - t_k|)$ and $\sigma_w^2 \min(t_{ij}, t_{ik})$. The $\epsilon_i$ are independently normally distributed with mean 0 and variance $\sigma^2 I$. We denote by $V_i$ the covariance matrix of $Y_i$:

$$V_i = Z_i B Z_i^T + \Sigma_i + \sigma^2 I$$

3. The Likelihood

Let $Y^o_i$ be the $n^o_i$-vector of completely observed outcomes for subject $i$, and $Y^c_i$ the $n^c_i$-vector of censored observations ($n_i = n^o_i + n^c_i$). After reordering, $X_i$, $Y_i$ and $V_i$ can be partitioned as: $Y_i^T = [Y^o_i, Y^c_i]$, $X_i^T = [X^o_i, X^c_i]$ and

$$V_i = \begin{bmatrix} V^o_i & V^{o \cdot c}_i \\ V^{c \cdot o}_i & V^c_i \end{bmatrix}.$$  

According to model (1), $Y^o_i$ has a multivariate Gaussian probability density function, $f^o_i$, with expectation $\mu^o_i = X^o_i \beta$ and covariance $V^o_i$. Let $c_i$ be the $n^c_i$-vector of censoring values for subject $i$ and $\theta$ the vector of parameters to be estimated. Then, the likelihood of the data takes the form:

$$L(\theta) = \prod_{i=1}^{N} f^o_i(y^o_i|\theta) P(Y^c_i < c_i|Y^o_i, \theta).$$

Using properties of the multivariate normal distribution, we find that the conditional distribution of $Y^c_i$ given $Y^o_i$ is Gaussian with expectation $\mu^{c|o}_i$ and covariance $V^{c|o}_i$ given by:

$$\mu^{c|o}_i = X^c_i \beta + V^{c \cdot o}_i V^o_i^{-1} [Y^o_i - \mu^o_i] \text{ and } V^{c|o}_i = V^c_i - V^{c \cdot o}_i V^o_i^{-1} V^{o \cdot c}_i T.$$  \hspace{1cm} (2)

If we denote by $\Phi^{c|o}_i$ the multivariate normal distribution function for $Y^c_i$ given $Y^o_i$, the likelihood may be rewritten:

$$L(\theta) = \prod_{i=1}^{N} f^o_i(y^o_i|\theta) \Phi^{c|o}_i(c_i|\theta)$$  \hspace{1cm} (3)
and the log-likelihood has the form:

\[ l(\theta) = \sum_{i=1}^{N} -n_i^o \frac{\log(2\pi)}{2} - \frac{1}{2} \log|V_i^o| - \frac{1}{2} \left( y_i^o - \mu_i^o \right)^T V_i^{o^{-1}} \left( y_i^o - \mu_i^o \right) \]

\[ + \log \int_{-\infty}^{C_1} \int_{-\infty}^{C_2} \ldots \int_{-\infty}^{C_{n_i}^o} (2\pi)^{-\frac{n_i^o}{2}} |V_{i|o}^o|^{-\frac{1}{2}} \exp \left[ -\frac{1}{2} \left( u - \mu_i^{c,o} \right)^T V_{i|o}^{c,o^{-1}} \left( u - \mu_i^{c,o} \right) \right] du \]

where \( u \) is a \( n_i^o \) vector.

### 4. Computation

We developed a FORTRAN program to estimate model parameters by direct maximization of the likelihood (4) using an iterative process.

Computation of the likelihood requires numerical computation of the integral of a multivariate normal density for each subject with censored measures. This was performed by using a FORTRAN routine named SADMVN developed by Genz (1992). The algorithm transforms the integral into an integral over a unit hyper-cube, reorders the integration variables and then applies a subregion adaptive multiple integration method (Berntsen et al., 1991). This algorithm was found to be the best of five methods considered when the size of the integral is less than 10, which is the most frequent case for epidemiological applications (Genz, 1993).

In order to impose a positivity constraint for the covariance parameters, the likelihood was reparameterized in terms of the square root for \( \alpha \) and the logarithmic transformation for the autocorrelation coefficient \( \sigma \), and the Cholesky factorization for the covariance matrix of the random effects \( B \) (\( B = U^T U \) where \( U \) is an upper triangular matrix).

To reduce computation time for models with many covariance parameters, we combined two optimization algorithms. The first iterations were performed using the simplex algorithm, after which the Marquardt algorithm (Marquardt, 1963) was used near the optimum. However, the simulation study was performed using only the Marquardt algorithm. Starting values for the global optimization process were obtained from a mixed effects linear model by imputing the censoring threshold to censored measures.

The estimate of the variance–covariance matrix of fixed effects estimates was minus the inverse of the matrix of second derivatives of the log-likelihood (4). The derivatives were computed numerically by finite difference.

Rather than direct maximization of the likelihood (4), Hughes proposed to use the EM algorithm. Considering \( \gamma_i \) and \( Y_i^c \) as missing data (since the model does not include \( w_i \)), the EM algorithm consists in iterating between computation of the expectations of complete data sufficient statistics and maximization of the likelihood of complete data (the likelihood assuming that \( \gamma_i \) and \( Y_i^c \) are observed) replacing the sufficient statistics by their expectations. However, the E-step requires evaluation of integrals which are intractable by classical methods and the expectations are computed by Gibbs sampling leading to an MCEM algorithm.

### 5. Estimation of Subject-specific Deviations

When data are not censored, the subject-specific deviations from the linear mixed-effects model (1), \( W_i(u) = Z_i(u)\gamma_i + w_i(u) \), for a vector of time \( u \) (not necessarily observed) are estimated by empirical Bayes, ie by \( E(W_i(u)|Y_i(t_i), \hat{\theta}) \). The vector \( [W_i(u)Y_i(t_i)]^T \) has a multivariate normal (MVN) distribution:

\[
\begin{bmatrix} W_i(u) \\ Y_i(t_i) \end{bmatrix} \sim \text{MVN} \left( \begin{bmatrix} 0 \\ X_i(t_i)\beta \end{bmatrix}, \begin{bmatrix} Z_i(u)BZ_i^T(u) + \Sigma_i(u) & Z_i(u)BZ_i^T(t_i) + \Sigma_i(u, t_i) \\ Z_i(t_i)BZ_i^T(u) + \Sigma_i(t_i, u) & \Sigma_i(t_i) \end{bmatrix} \right)
\]
where \( \Sigma_i(u, t_i) \) is the covariance between the vectors \( w_i(u) \) and \( w_i(t_i) \). Thus:

\[
E \{ W_i(u) | Y_i(t_i) \} = \left[ Z_i(u) B Z_i^t(t_i) + \Sigma_i(u, t_i) \right] V_i(t_i)^{-1} \left[ Y_i(t_i) - X_i(t_i) \hat{\beta} \right]. \tag{5}
\]

In the censored case, the expectation of \( W_i(u) \) given the data is

\[
E \{ W_i(u) | \text{data} \} = E \{ E \{ W_i(u) | Y_i(t_i) \} | Y_i^o = y_i^o, Y_i^c < c_i \}
\]

and using (5) we obtain:

\[
E \{ W_i(u) | \text{data} \} = E \left[ E \{ W_i(u) | Y_i(t_i) \} \right] V_i(t_i)^{-1} \left[ E \left\{ Y_i(t_i) | Y_i^o = y_i^o, Y_i^c < c_i \right\} - X_i(t_i) \hat{\beta} \right]. \tag{6}
\]

In (6), \( E \left\{ Y_i^c(t_i) | Y_i^o = y_i^o, Y_i^c < c_i \right\} = y_i^c \), but the conditional expectation of \( Y_i^c \) is intractable and was therefore estimated by simulation. Specifically, \( E \left\{ Y_i^c(t_i) | Y_i^o = y_i^o, Y_i^c < c_i \right\} \) was estimated by the empirical mean of 1000 vectors \( S \) of size \( n_i^c \) simulated according to the multivariate Gaussian distribution \( f_{Y_i^c | Y_i^o} \) defined by (2) and such that \( S < c_i \) (using a rejection method). A probably more rapid alternative would be to compute this expectation by Gibbs sampling as done by Hughes (1999). However, in our approach, this expectation is computed only once, after convergence, to estimate subject-specific deviations or random effects. Thus, the computation time is not a problem. Finally, an estimate of \( W_i(u) \) is obtained by replacing the vector \( \beta \) and the covariance parameters by their maximum likelihood estimators.

When \( u = t_i \), formula (6) becomes:

\[
E \{ W_i(t_i) | \text{data} \} = [I - \sigma^2 V_i(t_i)^{-1}] \left[ E \{ Y_i(t_i) | Y_i^o = y_i^o, Y_i^c < c_i \} - X_i(t_i) \hat{\beta} \right]. \tag{7}
\]

With the same approach, the random effects may be estimated by empirical Bayes, \( \hat{\gamma}_i = B Z_i(t_i) V_i(t_i)^{-1} (Y_i(t_i) - X_i(t_i) \hat{\beta}) \), replacing \( Y_i(t_i) \) by \( E \left\{ Y_i(t_i) | Y_i^o = y_i^o, Y_i^c < c_i \right\} \).

Using the estimates of \( W_i(u) \) defined by (6) or (7), estimates of the individual outcomes may be computed by \( X_i(t_i) \hat{\beta} + \tilde{W}_i(t_i) = \hat{Y}_i(t_i) = \max(Y_i(t_i), c_i) \). These censored residuals may be used to check the assumed normal distribution of the independent error \( \epsilon_i \). Indeed, the cumulative frequency distribution of censored residuals may be computed by the Kaplan–Meier approach and compared to the cumulative distribution of a normally distributed variable with variance \( \sigma^2 \) (Marschner et al., 1999). Practically, the Kaplan–Meier estimator is computed by replacing the survival time by the value of the residuals and the censoring indicator indicates if the residual is right-censored or completely observed. A residual is right-censored if the measure \( Y_{ij} \) is below the detection limit whatever the predicted value. The Kaplan–Meier method estimates \( \Pr(\hat{Y}_{ij} - Y_{ij} < r) \) for all possible values of the residuals \( r \).

6. Simulation study

The aim of the simulation study was to compare estimates of parameters of the linear mixed effects model obtained by imputing the quantification limit to censored values, by the MCEM algorithm or by maximizing the likelihood of censored data by the Marquardt algorithm as described in Section 3 and 4. For the MCEM algorithm, we used the FORTRAN program implemented by Hughes (1999).

For comparisons, we used in both algorithms the convergence criterion used by Hughes, that is a maximum relative change on the parameters less than 0.01. However, in the Marquardt algorithm, two additional convergence criteria based on the gradient and on the change in the log-likelihood had to be satisfied to assume that convergence was reached. The optimization processes were stopped if they did not converge in 600 cpu seconds with a pentium II 300 Mhz PC (LINIX operating system). The Marquardt algorithm converged in less than 600 cpu seconds for 100% of the simulated samples. Convergence
problems with the MCEM algorithms are detailed below. Biases and mean squared errors (MSE) were computed on simulations which converged.

Data were simulated according to three linear mixed effects model with different covariance structures. The first model was the following random intercept model:

\[ Y_{ij} = \beta_0 + \beta_1 t_{ij} + \gamma_0 + \epsilon_{ij}. \]  
(8)

The second model included random slope and intercept and an independent error:

\[ Y_{ij} = \beta_0 + \beta_1 t_{ij} + \gamma_0 + \gamma_1 t_{ij} + \epsilon_{ij}. \]  
(9)

The third model included an autoregressive error and an independent error:

\[ Y_{ij} = \beta_0 + \beta_1 t_{ij} + w_i(t_{ij}) + \epsilon_{ij}. \]  
(10)

where \( \text{cov}(w_i(t_{ij}), w_i(t_{ik})) = \sigma^2_w \exp(-\alpha|t_{ij} - t_{ik}|) \). This model was not estimated by the Hughes’ MCEM algorithm because it does not accept autoregressive error.

Values of the parameters were chosen to be close to those obtained from the Aquitaine cohort of HIV-infected patients (Thiébaut et al., 2000) and receiving highly active anti-retroviral therapy. The parameters were \( \beta_0 = 4, \beta_1 = -0.5 \) and \( \sigma^2_{\epsilon} = 1 \). For model (8), \( \sigma^2_{\gamma_0} = 0.25 \), and for model (10) \( \sigma^2_w = 1 \) and \( \alpha = 0.1 \). For model (9), the covariance matrix for the random effects was:

\[ B = \begin{pmatrix} 0.25 & -0.01 \\ -0.01 & 0.04 \end{pmatrix} \]

The censoring value was independent of time and subject (\( c_{ij} = c \forall i, j \)). We simulated samples with approximately 17% (\( c = 1 \)) and 36% (\( c = 2 \)) of censored measures. For each model and each censoring threshold, 500 samples of 100 subjects were simulated. The number of repeated measures for each subject was randomly distributed between 2 and 7 (mean 4) and the times of measurements were uniformly distributed between 0 and 6.

Results are presented in Tables 1 to 3. For the three models, the bias and the MSE of the fixed parameters were much smaller when the censoring was taken into account, whatever the algorithm. The intercept was less biased than the slope with the crude imputation method because very few measures were below the censoring value for times close to 0 (representing the time of therapy initiation). The bias and MSE of the covariance parameters were also reduced when censoring was taken into account with the Marquardt algorithm compared to the crude imputation method.

When the model included only a random intercept, the bias and the MSE were similar with the MCEM and Marquardt algorithms (Table 1). The computation time depended on the number of measures \( n_i \) for each subject. The Marquardt algorithm was more rapid for small \( n_i \) (\( n_i \) between 2 and 7 with mean 4 for instance) and slower for large \( n_i \) (\( n_i = 7 \forall i \) for instance; results not displayed here). When \( c = 2 \), 15% (for \( 2 \leq n_i \leq 7 \)) or 5% (for \( n_i = 7 \forall i \)) of the MCEM optimizations did not converge.

When the model included two random effects, the bias and the MSE were larger with the MCEM algorithm, particularly for covariance parameters (Table 2). Moreover, the Marquardt algorithm always converged in less than 600 seconds (mean time < 60s), while 39% (when \( c = 1 \)) and 26% (when \( c = 2 \)) of the MCEM optimizations did not converge. For successful MCEM optimizations, the mean cpu time until convergence was 71 (when \( c = 1 \)) and 43 seconds (when \( c = 2 \)). Results of the MCEM algorithm were not improved with larger \( n_i \) (\( n_i = 7 \forall i \)) of the MCEM optimizations did not converge. For successful MCEM optimizations, the mean cpu time until convergence was 71 (when \( c = 1 \)) and 43 seconds (when \( c = 2 \)). Results of the MCEM algorithm were not improved with larger \( n_i \) (\( n_i = 7 \forall i \)). However, additional simulations showed that the behaviour of the MCEM algorithm depended highly on the ratio between the variance explained by random effects and the variance of the independent error. With the same values for the covariance matrix B, the proportion of successful convergences increased and the biases and MSE of the estimates decreased.
Table 1. Bias and mean squared error (MSE) for the parameters of the mixed effects linear model \( Y_{ij} = \beta_0 + \beta_1 t_{ij} + \gamma_{0i} + \epsilon_{ij} \) estimated by imputing the quantification limit to censored measures, by the MCEM algorithm or by the Marquardt algorithm (500 replications)

<table>
<thead>
<tr>
<th>% of censoring</th>
<th>Method</th>
<th>Imputation of the limit</th>
<th>Marquardt algorithm</th>
<th>MCEM algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bias (%)</td>
<td>MSE</td>
<td>Bias (%)</td>
</tr>
<tr>
<td>17%</td>
<td>( \beta_0 )</td>
<td>-2.72</td>
<td>0.026</td>
<td>-0.35</td>
</tr>
<tr>
<td></td>
<td>( \beta_1 )</td>
<td>-13.8</td>
<td>0.006</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_0 )</td>
<td>-29.9</td>
<td>0.008</td>
<td>-2.44</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-17.2</td>
<td>0.032</td>
<td>5.50</td>
</tr>
<tr>
<td>36%</td>
<td>( \beta_0 )</td>
<td>-4.69</td>
<td>0.048</td>
<td>-0.94</td>
</tr>
<tr>
<td></td>
<td>( \beta_1 )</td>
<td>-36.2</td>
<td>0.033</td>
<td>-2.52</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_0 )</td>
<td>-60.5</td>
<td>0.024</td>
<td>-4.36</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-43.9</td>
<td>0.195</td>
<td>3.65</td>
</tr>
</tbody>
</table>

Table 2. Bias and mean squared error (MSE) for the parameters of the mixed effects linear model \( Y_{ij} = \beta_0 + \beta_1 t_{ij} + \gamma_{0i} + \gamma_{1i} t_{ij} + \epsilon_{ij} \) estimated by imputing the quantification limit to censored measures, by the MCEM algorithm or by the Marquardt algorithm (500 replications)

<table>
<thead>
<tr>
<th>% of censoring</th>
<th>Method</th>
<th>Imputation of the limit</th>
<th>Marquardt algorithm</th>
<th>MCEM algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bias (%)</td>
<td>MSE</td>
<td>Bias (%)</td>
</tr>
<tr>
<td>17%</td>
<td>( \beta_0 )</td>
<td>-3.80</td>
<td>0.038</td>
<td>-0.44</td>
</tr>
<tr>
<td></td>
<td>( \beta_1 )</td>
<td>-20.0</td>
<td>0.011</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_0 )</td>
<td>63.7</td>
<td>0.054</td>
<td>-0.79</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{01} )</td>
<td>197.2</td>
<td>0.002</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-71.0</td>
<td>0.0009</td>
<td>-0.32</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-17.7</td>
<td>0.035</td>
<td>2.95</td>
</tr>
<tr>
<td>36%</td>
<td>( \beta_0 )</td>
<td>-5.82</td>
<td>0.067</td>
<td>-0.91</td>
</tr>
<tr>
<td></td>
<td>( \beta_1 )</td>
<td>-42.9</td>
<td>0.047</td>
<td>-2.50</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_0 )</td>
<td>104.0</td>
<td>0.089</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>( \sigma_{01} )</td>
<td>611.8</td>
<td>0.005</td>
<td>70.1</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-70.6</td>
<td>0.0008</td>
<td>-5.52</td>
</tr>
<tr>
<td></td>
<td>( \sigma^2_\epsilon )</td>
<td>-42.7</td>
<td>0.185</td>
<td>4.64</td>
</tr>
</tbody>
</table>
Table 3. Bias and mean squared error (MSE) for the parameters of the mixed effects linear model $Y_{ij} = \beta_0 + \beta_1 t_{ij} + w_i(t_{ij}) + \epsilon_{ij}$ estimated by imputing the quantification limit to censored measures or by the Marquardt algorithm (500 replications)

<table>
<thead>
<tr>
<th>% of censoring</th>
<th>Parameter</th>
<th>Method of the limit</th>
<th>Bias (%)</th>
<th>MSE</th>
<th>Bias (%)</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>Imputation</td>
<td>-2.46</td>
<td>0.036</td>
<td>-0.39</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td></td>
<td>-18.1</td>
<td>0.009</td>
<td>-0.57</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_w$</td>
<td></td>
<td>-27.2</td>
<td>0.092</td>
<td>3.21</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td></td>
<td>19.2</td>
<td>0.005</td>
<td>-7.13</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_\epsilon$</td>
<td></td>
<td>-23.7</td>
<td>0.064</td>
<td>4.06</td>
<td>0.015</td>
</tr>
<tr>
<td>36%</td>
<td>$\beta_0$</td>
<td>Marquardt</td>
<td>-3.06</td>
<td>0.037</td>
<td>-0.74</td>
<td>0.033</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$</td>
<td></td>
<td>-38.5</td>
<td>0.038</td>
<td>-1.54</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_w$</td>
<td></td>
<td>-51.9</td>
<td>0.279</td>
<td>3.76</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>$\alpha$</td>
<td></td>
<td>77.7</td>
<td>0.016</td>
<td>-6.55</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_\epsilon$</td>
<td></td>
<td>-48.3</td>
<td>0.240</td>
<td>2.39</td>
<td>0.018</td>
</tr>
</tbody>
</table>

when $\sigma^2_\epsilon$ decreased. For instance, when $\sigma^2_\epsilon = 0.1$ and $c = 2$, 92% of the MCEM optimizations converged and the biases were similar with both methods (biases were less than 2% for all parameters except for $\sigma_{01}$ which was 8%). The bad results of the MCEM algorithm for some covariance structures may be explained by the slow convergence of the EM algorithm and by the unavailability of a convergence criterion based on the derivatives of the log-likelihood, which was already described as one of the major drawbacks of the EM algorithm (Lindstrom and Bates, 1988). Indeed, the convergence criterion of the MCEM algorithm is only based on the relative change in the parameters and, with the Marquardt algorithm, this criterion was often reached long before the criterion based on the derivatives.

7. APPLICATION

The simulation study showed the usefulness of this method when model assumptions are verified. We have also applied the method to a real data set. The data came from the randomized clinical trial ALBIBANRS 070 which compared three anti-retroviral treatment regimens in HIV-1-infected patients: stavudine (d4T) plus didanosine (ddI) for 24 weeks, zidovudine (AZT) plus lamivudine (3TC) for 24 weeks and a switching group with the d4T + ddI regimen for 12 weeks followed by the AZT + 3TC regimen for 12 weeks. The patients included in the trial were adults (more than 18 years old), naive of anti-retroviral therapy, with baseline CD4+ count of 200 cells/$\mu l$ or more and a baseline plasma HIV-1 RNA level between 10 000 and 100 000 copies/ml. HIV-1 RNA level was measured at baseline and then approximately every 4 weeks until week 24 with two assays: a classical branched DNA assay (Quantiplex HIV-1 RNA 2.0, Chiron Diagnostics) with a lower limit of quantification of 500 copies/ml and an ultrasensitive polymerase chain reaction (PCR) assay with a lower limit of quantification of 50 copies/ml (Ultrasensitive
Fig. 1. Mean RNA levels measured with the ultrasensitive assay at the seven visits and by replacing the censored measures by the quantification limit of 50 copies/ml (ALBI ANRS 070); long dashed line: group AZT + 3TC, plain line: switching group, short dashed line: group d4T + ddI, vertical bars with tick-marks: 95% confidence intervals.

amplicor HIV-1 1.5 Monitor, Roche Molecular Systems). The method and the results of this trial using the ultrasensitive assay were detailed by Molina et al. (1999).

This re-analysis of the ALBI data aimed to compare the results regarding the evolution of HIV-1 RNA level and the effect of treatments, according to the statistical method and the quantification limit of the measurements. However, comparisons of results based on data obtained with different assays would be blurred because the differences between the assays may be large (Holguin et al., 1999). So, we did not use the bDNA data set but created another data set by using the ultrasensitive measures and artificially censoring them below 500 copies/ml, leading to a percentage of censored values of 50% (PCR 500 data set). In the original data set (PCR 50), the percentage of censored measures was 17%. For this analysis, the sample included 148 patients which had at least two measures of RNA level: 49 subjects in both AZT + 3TC and switching groups, and 50 in the d4T + ddI group.

Figure 1 shows the mean HIV-1 RNA levels in log_{10} scale as measured by the ultrasensitive assay (PCR 50) at the seven visits when replacing the censored measures by the quantification limit (50 copies/ml = 1.7 log_{10} copies/ml). This suggests a sharp decrease in the viral load in the three groups during the first month. Then the viral load remains around 2 log_{10} copies/ml for the d4T + ddI group and increases slowly for the AZT + 3TC group. The HIV-1 RNA level in the switching group evolved in the same way as the d4T + ddI until week 16 (the first visit after switching treatment) but then increased until week 24.

To take this evolution into account, we fitted a model with an intercept, a slope for the first month and another slope for the end of the study, in the fixed and random components. The date of change of the slope was estimated to be \( t_1 = 23 \) days using a profile likelihood with the PCR 50 data set by assigning the quantification limit to censored measures. The first model estimated on the two data sets for the log_{10} of the RNA level was the following (model 1):

\[
Y_{ij} = \beta_0 + \beta_1 \min(t_{ij}, t_1) + \beta_2 (t_{ij} - t_1)I_{t_{ij} > t_1} + \gamma_0 + \gamma_1 \min(t_{ij}, t_1) + \gamma_2 (t_{ij} - t_1)I_{t_{ij} > t_1} + \epsilon_{ij}
\]

where \( \gamma_i = [\gamma_{0i}, \gamma_{1i}, \gamma_{2i}]^T \) is trivariate normal with covariance matrix \( B \), of which the diagonal elements
are denoted by $\sigma_0^2$, $\sigma_1^2$. Including treatment effects, we estimated model 2:

$$Y_{ij} = \beta_0 + \beta_1 \min(t_{ij}, t_1) + \beta_2(t_{ij} - t_1)I_{ij} > t_1 + \gamma_0i + \gamma_1i \min(t_{ij}, t_1) + \gamma_2(t_{ij} - t_1)I_{ij} > t_1 + \epsilon_{ij} + \beta_{01}I_{d4T+ddl} + \beta_{11} \min(t_{ij}, t_1)I_{d4T+ddl} + \beta_{21}(t_{ij} - t_1)I_{ij} > t_1I_{d4T+ddl} + \beta_{02}I_{switch} + \beta_{12} \min(t_{ij}, t_1)I_{switch} + \beta_{22}(t_{ij} - t_1)I_{ij} > t_1I_{switch}$$

where $I_{d4T+ddl}$ and $I_{switch}$ are indicator variables for groups $d4T + ddf$ and switching, respectively, group $AZT + 3TC$ being considered as the reference group.

Parameters were estimated by imputing the quantification limit or half the limit, by the MCEM algorithm and by the proposed method (SM, combining Simplex and Marquardt algorithms). For the comparisons, parameters estimated by taking censoring into account using the PCR 50 data set were considered as the gold standard.

Results for model 1 are shown in Table 4. Whatever the data set and the method, the mean viral load exhibited a significant sharp decrease during the first 23 days and a significant slow increase after that time. However, in the data set censored at 500 copies/ml ($= 2.7 \log_{10}$), when the quantification limit or half the limit were imputed, the first slope was underestimated. Consequently, with the PCR 500 data set, the estimated mean HIV-1 RNA level at 23 days was 2.83 $\log_{10}$ copies/ml by imputing the quantification limit, 2.61 by imputing half the limit, 2.21 by the MCEM algorithm and 2.25 by the SM algorithm. With the PCR 50 data set, these values were respectively 2.43, 2.40, 2.41 and 2.41 $\log_{10}$ copies/ml. As expected, the estimates of the variances of the fixed parameters and the estimates of the variance components (except $\sigma_1^2$) were biased downwards with the crude imputation methods. When the censoring was taken into account, the MCEM and SM estimates from the two data sets were close. However, for the PCR 500 data set, both methods overestimated the slope for the second period of time and the variance of the first random slope. The computing times were 75 s (SM) and 4 s (MCEM) for the PCR 50 data set and 14 min (SM) and 2 hours and 40 min (MCEM) for the PCR 500 data set.

Results for model 2 are displayed in Table 5. As the main objective was to compare results regarding treatment effects, this model was estimated by imputing only the quantification limit (the conservative strategy) and not half the limit. Moreover, for the PCR 500 data set, the MCEM algorithm did not reach the convergence criterion after 2 days of computation. For the PCR 50 data set, the MCEM and SM estimates were similar.

For the group $AZT + 3TC$ (the reference), both slopes were significantly different from 0, whatever the method and the data set, but the bias was very large when the censoring threshold was 500 copies/ml with the crude imputation method. With the SM algorithm, the slopes estimated from the two data sets were very close.

In the $d4T + ddl$ and switching groups, the first slope was not significantly different from that of the reference group, whatever the method and the data set. On the contrary, the slope after the first 23 days for these groups was significantly different from that of the reference group except with the SM algorithm and the PCR 500 data set. Compared to our ‘gold standard’, the estimates of the slopes obtained with the crude imputation method were biased when the censoring threshold was 500 copies/ml. For instance, the estimates of the first slope was $-7.27$ compared to $-9.13$ for the $d4T + ddl$ group and $-7.65$ compared to $-9.67$ for the switching group. However, for these two groups, the estimates of the slopes were different when the censoring threshold changed even if the censoring was taken into account. Moreover, the standard errors of these estimates were large. In fact, estimates of the slopes were difficult in these groups with the PCR 500 data set because the percentages of censored measures after the baseline visit were very high: 79% for the $d4T + ddl$ group and 64% for the switching group versus 35% for the $AZT + 3TC$ group.

The right-censored residuals adjusted for the treatment group (model 2) were computed for the PCR 50 data set. Figure 2 displays the Kaplan–Meier estimate of the cumulative frequency distribution of
Table 4. Estimates of the mixed effects linear model (model 1) for the log_{10} of the HIV-1 RNA level according to the data sets and the methods for dealing with censoring (ALBI-ANRS 070 trial). The slopes are estimated for 100 days

<table>
<thead>
<tr>
<th>Data set and method</th>
<th>$\hat{\beta}_0$ (sd)</th>
<th>$\hat{\beta}_1$ (sd)</th>
<th>$\hat{\beta}_2$ (sd)</th>
<th>$\sigma^2_0$</th>
<th>$\sigma^2_1$</th>
<th>$\sigma^2_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PCR 50</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantification limit</td>
<td>4.53 (0.03)</td>
<td>−9.13 (0.25)</td>
<td>0.23 (0.05)</td>
<td>4.22 (0.30)</td>
<td>0.15 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Half the limit</td>
<td>4.54 (0.04)</td>
<td>−9.30 (0.27)</td>
<td>0.21 (0.06)</td>
<td>4.86 (0.35)</td>
<td>0.18 (0.06)</td>
<td></td>
</tr>
<tr>
<td>MCEM</td>
<td>4.54 (0.04)</td>
<td>−9.26 (0.28)</td>
<td>0.16 (0.07)</td>
<td>5.01 (0.47)</td>
<td>0.19 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Simplex/Marquardt</td>
<td>4.54 (0.04)</td>
<td>−9.28 (0.28)</td>
<td>0.17 (0.06)</td>
<td>5.04 (0.44)</td>
<td>0.19 (0.06)</td>
<td></td>
</tr>
<tr>
<td><strong>PCR 500</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantification limit</td>
<td>4.53 (0.03)</td>
<td>−7.40 (0.19)</td>
<td>0.20 (0.04)</td>
<td>3.24 (0.13)</td>
<td>0.068 (0.04)</td>
<td></td>
</tr>
<tr>
<td>Half the limit</td>
<td>4.53 (0.03)</td>
<td>−8.34 (0.34)</td>
<td>0.26 (0.07)</td>
<td>3.26 (0.18)</td>
<td>0.27 (0.07)</td>
<td></td>
</tr>
<tr>
<td>MCEM</td>
<td>4.53 (0.04)</td>
<td>−10.1 (0.38)</td>
<td>0.37 (0.07)</td>
<td>10.9 (0.41)</td>
<td>0.16 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Simplex/Marquardt</td>
<td>4.53 (0.03)</td>
<td>−9.92 (0.38)</td>
<td>0.36 (0.08)</td>
<td>9.09 (0.38)</td>
<td>0.16 (0.08)</td>
<td></td>
</tr>
</tbody>
</table>

these residuals and the cumulative distribution of a normal variable with mean zero and variance equal to the estimated variance of the independent error in model 2. This suggests that the distribution of the independent error is close to the assumed Gaussian distribution.

8. DISCUSSION

We have proposed a maximum likelihood approach to estimate mixed effects linear models with left-censored longitudinal data. The simulation study showed that the bias and mean squared error of parameter estimates obtained by this method are smaller than those obtained by imputing the censoring limit or half the limit. As expected, the application showed that the estimates taking censoring into account are more robust to changes of the censoring threshold.

For the maximization of the likelihood, we use a combination of a Simplex algorithm and a Marquardt algorithm rather than the MCEM algorithm recently proposed by Hughes (1999). For a simple covariance structure, estimates from both methods have similar properties but, for some complex covariance structures, the MCEM algorithm has more convergence problems and leads to more biased estimates. In addition, the MCEM algorithm was not developed for models including autoregressive error.

Analysis of the ALBI sample also demonstrated the limitation of the proposed method when the in-
Table 5. Estimates of the mixed effects linear model including treatment effect (model 2) for the log10 of the HIV-1 RNA level according to the data sets and the methods for dealing with the censoring (ALBI-ANRS 070 trial). The slopes are estimated for 100 days

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Data set and method</th>
<th>PCR 50</th>
<th>PCR 500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μ̂0</td>
<td>μ̂1</td>
<td>μ̂2</td>
</tr>
<tr>
<td></td>
<td>(sd)</td>
<td>(sd)</td>
<td>(sd)</td>
</tr>
<tr>
<td>AZT + 3TC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(reference)</td>
<td>4.57*</td>
<td>-8.77*</td>
<td>0.58*</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.43)</td>
<td>(0.08)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>d4T + ddI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching</td>
<td>4.57*</td>
<td>-9.03*</td>
<td>0.61*</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.48)</td>
<td>(0.10)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>PCR 50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quantification limit</td>
<td>4.57*</td>
<td>-7.30*</td>
<td>0.39*</td>
</tr>
<tr>
<td>(0.05)</td>
<td>(0.33)</td>
<td>(0.06)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>MCEM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simplex/Marquardt</td>
<td>4.57*</td>
<td>-9.00*</td>
<td>0.62*</td>
</tr>
<tr>
<td>(0.06)</td>
<td>(0.57)</td>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

*P < 0.05

Fig. 2. Kaplan–Meier estimate of the cumulative frequency distribution of the residuals of model 2 for the PCR 50 data set (plain line) with its 95% confidence interval (thin lines) and cumulative Gaussian distribution with mean zero and variance equal to the estimated variance of the independent error in model 2 (dashed line).
formation is not sufficient, i.e. when the proportion of censored measures is very large and the censoring threshold is identical for all subjects and all times. For instance, in the ALBI clinical trial, it was not possible to obtain reliable estimates of the slopes for group d4T + ddI when the censoring threshold was 500 copies/ml.

Another limitation is that model assumptions are difficult to verify because classical methods for regression diagnosis are not available for censored residuals. However, the Kaplan–Meier method may be used to check normality of the residuals.

We did not deal with right-censoring which may be due to an upper limit of quantification of the assay. In fact, this is less frequent because an exact measure may be obtained by re-assaying the sample after dilution.

As a maximum likelihood approach, the proposed method requires missing data to be ignorable: the probability of response must depend neither on the missing measures nor on the parameters of the model for the outcome. This assumption was very likely in the ALBI trial because most of these missing measures were due to technical problems. Moreover, a bias was unlikely because only 7.7% of the measures were missing in that study.

In conclusion, the likelihood approach allows us to estimate parameters from mixed effects linear models with general covariance structures and to obtain individual predictions for longitudinal Gaussian data when data are left-censored. The proposed algorithm corrects the bias obtained when an arbitrary value is assigned to censored data and gives better results than the MCEM algorithm for models with complex covariance structures.

ACKNOWLEDGEMENTS

We thank the Scientific Committee of the ALBI trial for authorizing us to use their data, the management centre at INSERM Unit 330 for providing us with the data and the statistician Valérie Journot for valuable discussions.

REFERENCES


[Received October 4, 1999; first revision March 6, 2000; second revision April 3, 2000; accepted for publication April 6, 2000]